Terascale Computing in Accelerator Science & Technology

Robert D. Ryne

Accelerator Modeling and Advanced Computing Group

Lawrence Berkeley National Laboratory

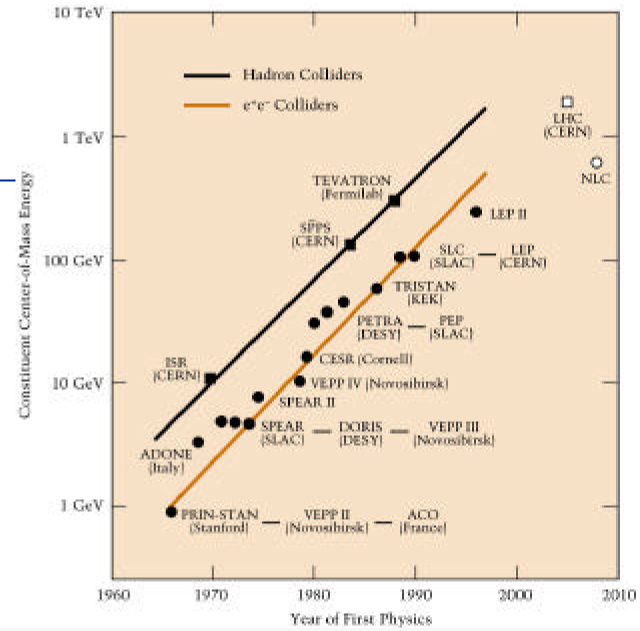
Acknowledgements:

Kwok Ko, David Bailey, Horst Simon,

the Computer Museum History Center, and many others

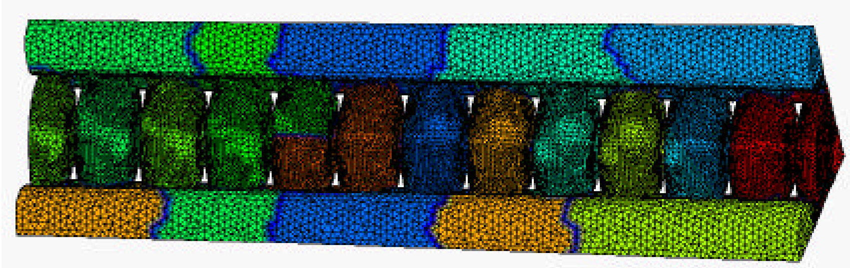
Outline

- Part I: Trends in Accelerators and High Performance Computing (HPC)
  —Livingston, Moore



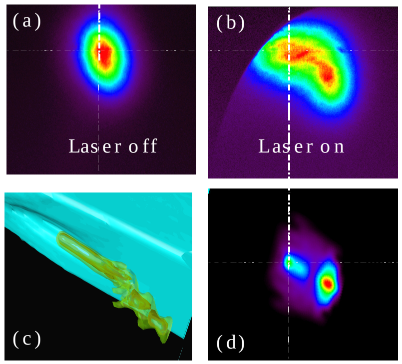
- Intermission

- Part II: Role of HPC in next-generation accelerator design



- Intermission

- Part III: Future challenges in HPC and accelerator development



The meaning of terascale

- Problem requirements
  - —trillions of floating point operations per sec (TFLOPS)
  - —trillions of bytes of memory (TBytes)
- Present-day example: IBM SP at NERSC
  - —3.75 TFLOPS, 1.7 TBytes
  - —158 "nodes" x 16 CPUs/node = 2528 CPUs

National Energy Research Scientific Computing Center (NERSC)

Motivation

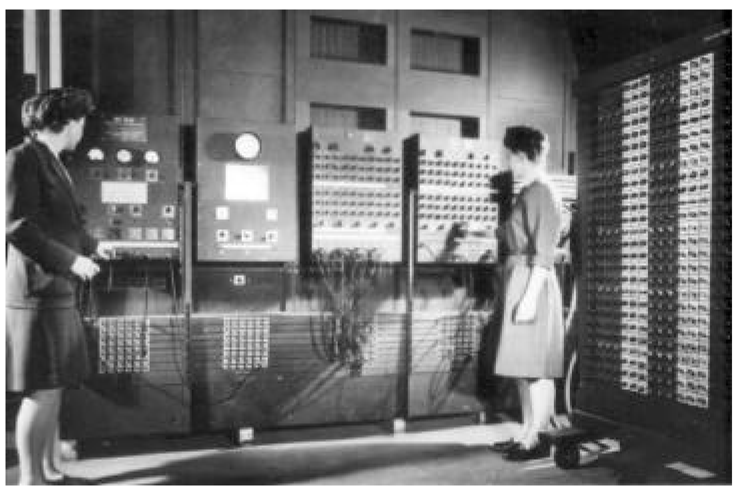
"...With the advent of everyday use of elaborate calculations, speed has become paramount to such a high degree that there is no machine on the market today capable of satisfying the full demand of modern computational methods. The most advanced machines have greatly reduced the time required for arriving at solutions to problems which might have required months or days by older procedures. This advance, however, is not adequate for many problems encountered in modern scientific work and the present invention is intended to reduce to seconds such lengthy computations..."

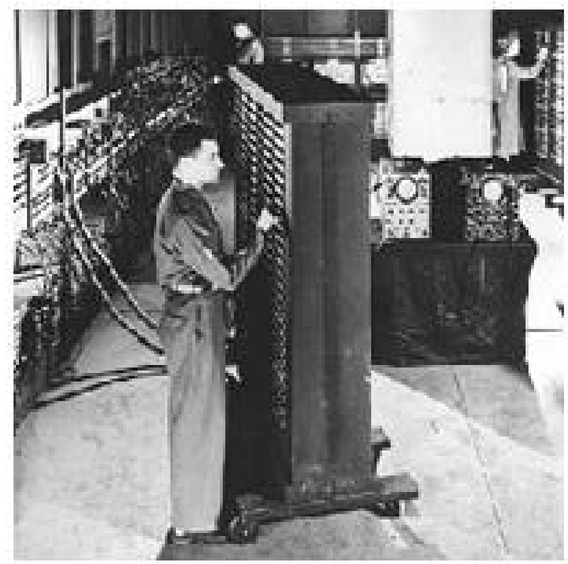
Motivation

"...With the advent of everyday use of elaborate calculations, speed has become paramount to such a high degree that there is no machine on the market today capable of satisfying the full demand of modern computational methods. The most advanced machines have greatly reduced the time required for arriving at solutions to problems which might have required months or days by older procedures. This advance, however, is not adequate for many problems encountered in modern scientific work and the present invention is intended to reduce to seconds such lengthy computations..."

From the ENIAC patent, 26 June 1947!

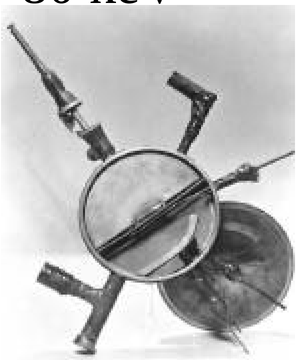
4 x 10^{-9} TFLOPS



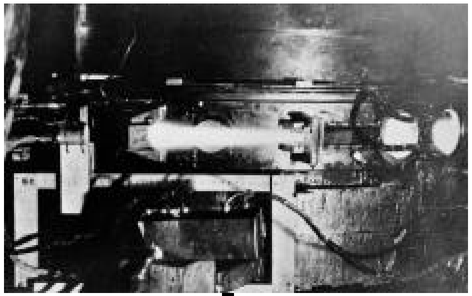


1930s

1st cyclotron
80 keV



27’’ cyclotron
4.8 MeV



60’’ cyclotron: 16 MeV

1930 — 1940

Wideroe linac: 1.2 MeV
1.14 m long tube

11’’ cyclotron
1.22 MeV

37’’ cyclotron
8 MeV

1940s

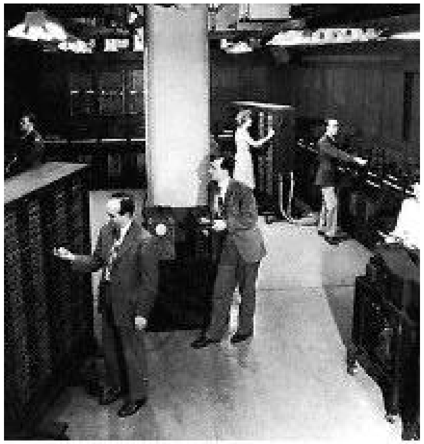
184’’ cyclotron
195 MeV

Alvarez linac
32 MeV, 40’

1940 — 1950

19000 tubes

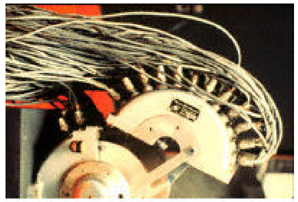
plug & socket
programs



ENIAC (4K adds/sec)

ESDAC (714 ops/sec)
1st stored program computer

Drum memory



1950s

Cosmotron (3 GeV)

Bevatron (6.2 GeV)

CERN PS (28 GeV)

Cornell
1.3 GeV

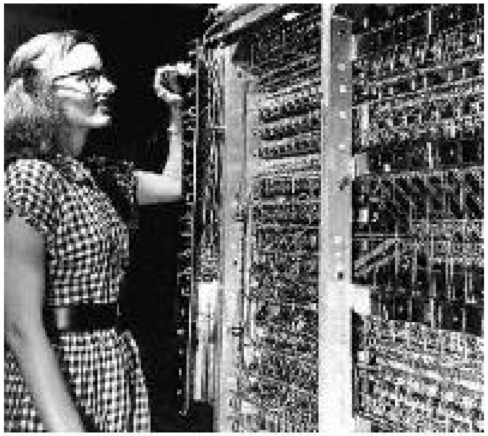
Antiprotons
detected

CERN
Synchro-
cyclotron
600 MeV

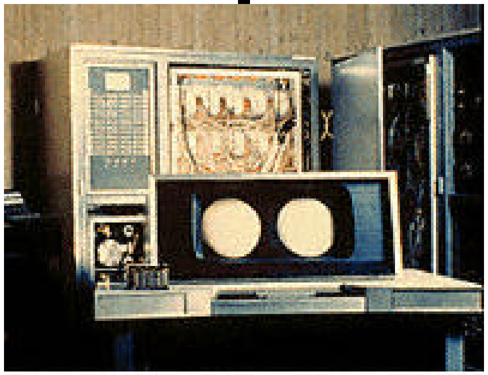
strong focusing

1950

1960



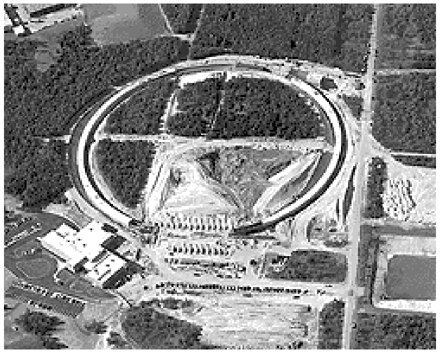
Von Neumann IAS



IBM's first
transistorized computer

1960s

Brookhaven
AGS: 33 GeV



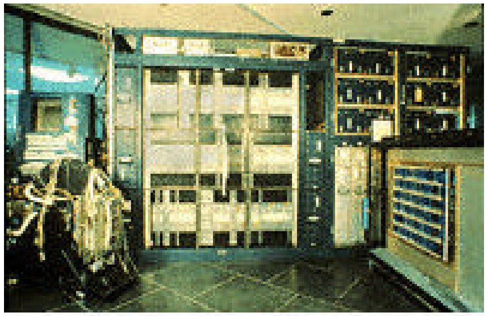
SLAC 2 mile
linac: 20 GeV

1960 — 1970

IBM 360

IBM 1401: transistors,
magnetic core memory

CDC 6600: 3 MIPS



ILLIAC IV: 300 MIPS

1970s

CERN ISR
(1st proton collider)

SPEAR,DORIS,VEPP III

Fermilab (500 GeV)

J/Psi

CERN SPS: 500 GeV

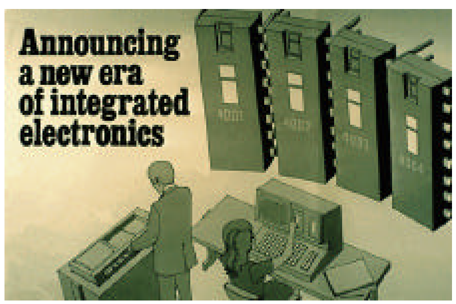
Stochastic
cooling

CESR

1970 1980

CDC 7600

Vector processors



Microprocessors
introduced

Cray 1
166 MFLOPS

1980s

1980

PETRA
PEP

SPS p-pbar
(100 GeV)

TRISTAN

FNAL Tevatron (2 TeV)

LEP
200 GeV

SLC
50 GeV

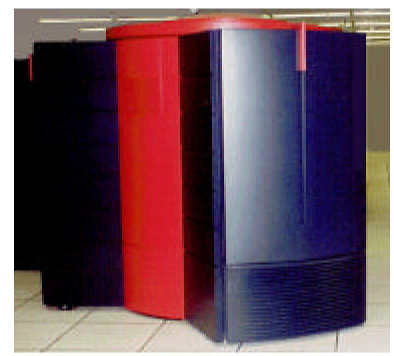
HERA

1990

Massively Parallel Processors

Cray XMP
477 MFLOPS

Connection Machine
10 GFLOPS



Cray C90

1990s

PEP-II, KEKB,
RHIC

Cancellation
of SSC

1990 — 2000

MPPs

Shared Memory (SMPs)

SMP Clusters

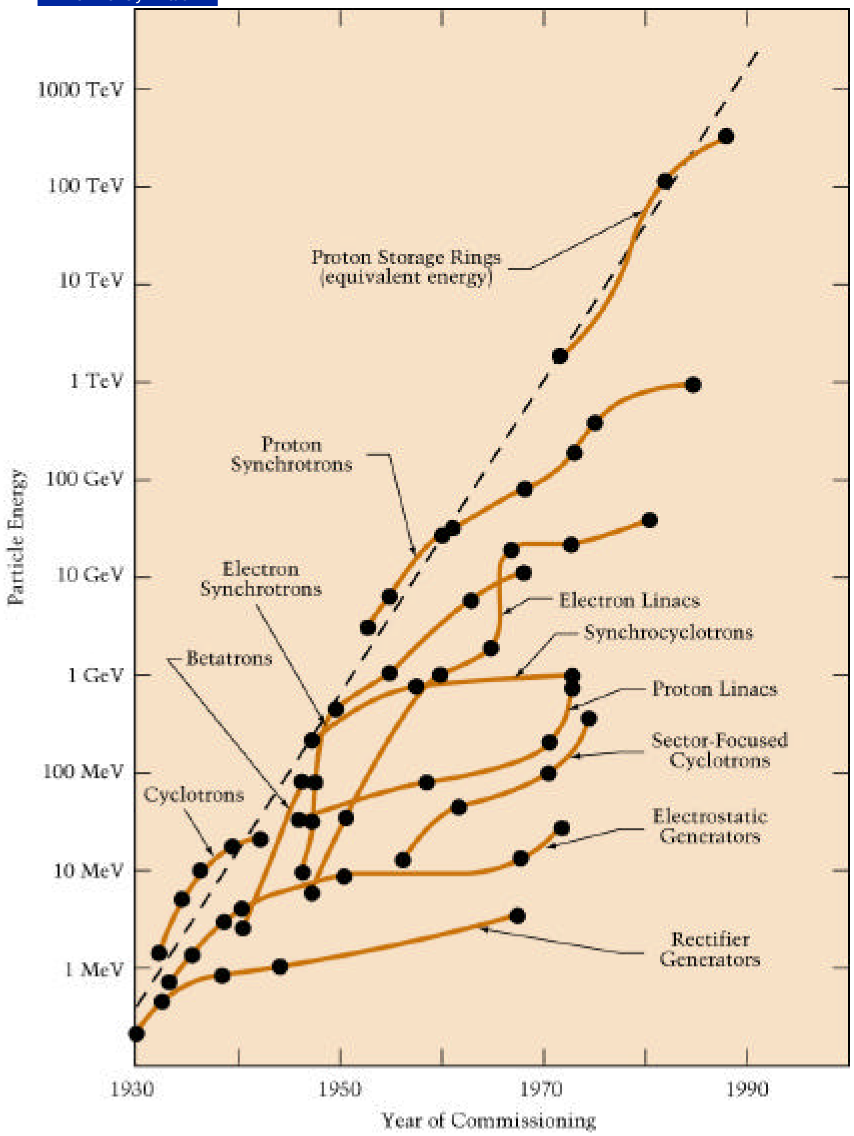
CM-5, Cray T3D
100 GFLOPS

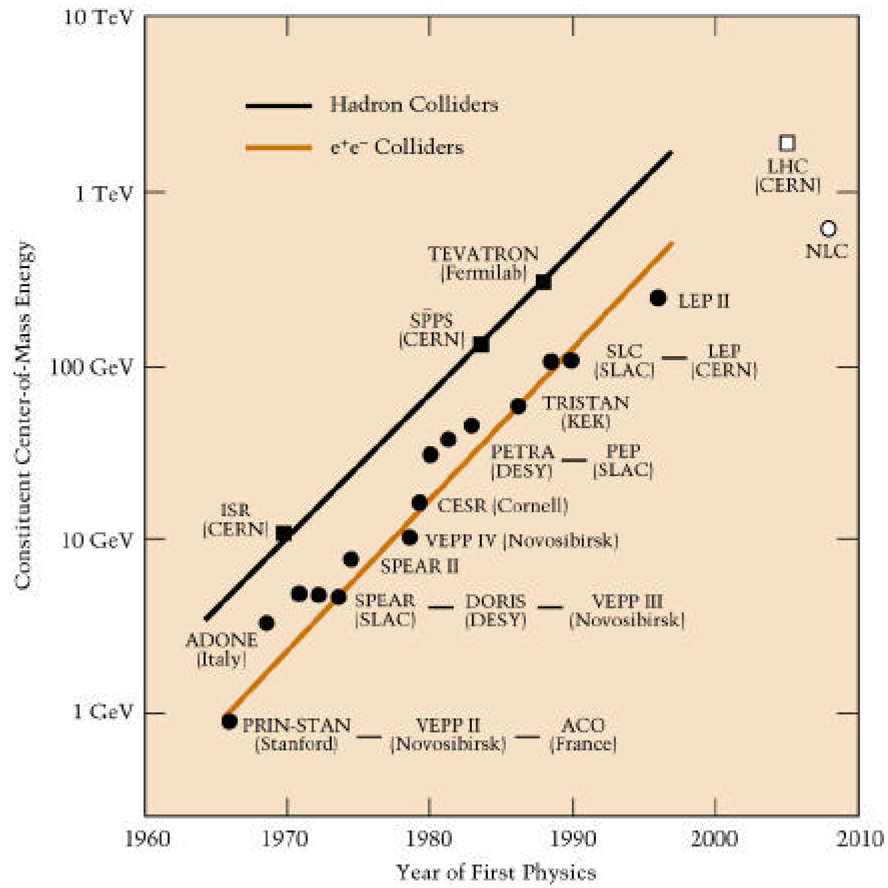
Cray T3E
450 GFLOPS
1st TFLOP appl.

ASCI Red
1 TFLOPS

ASCI Blue,
3 TFLOPS

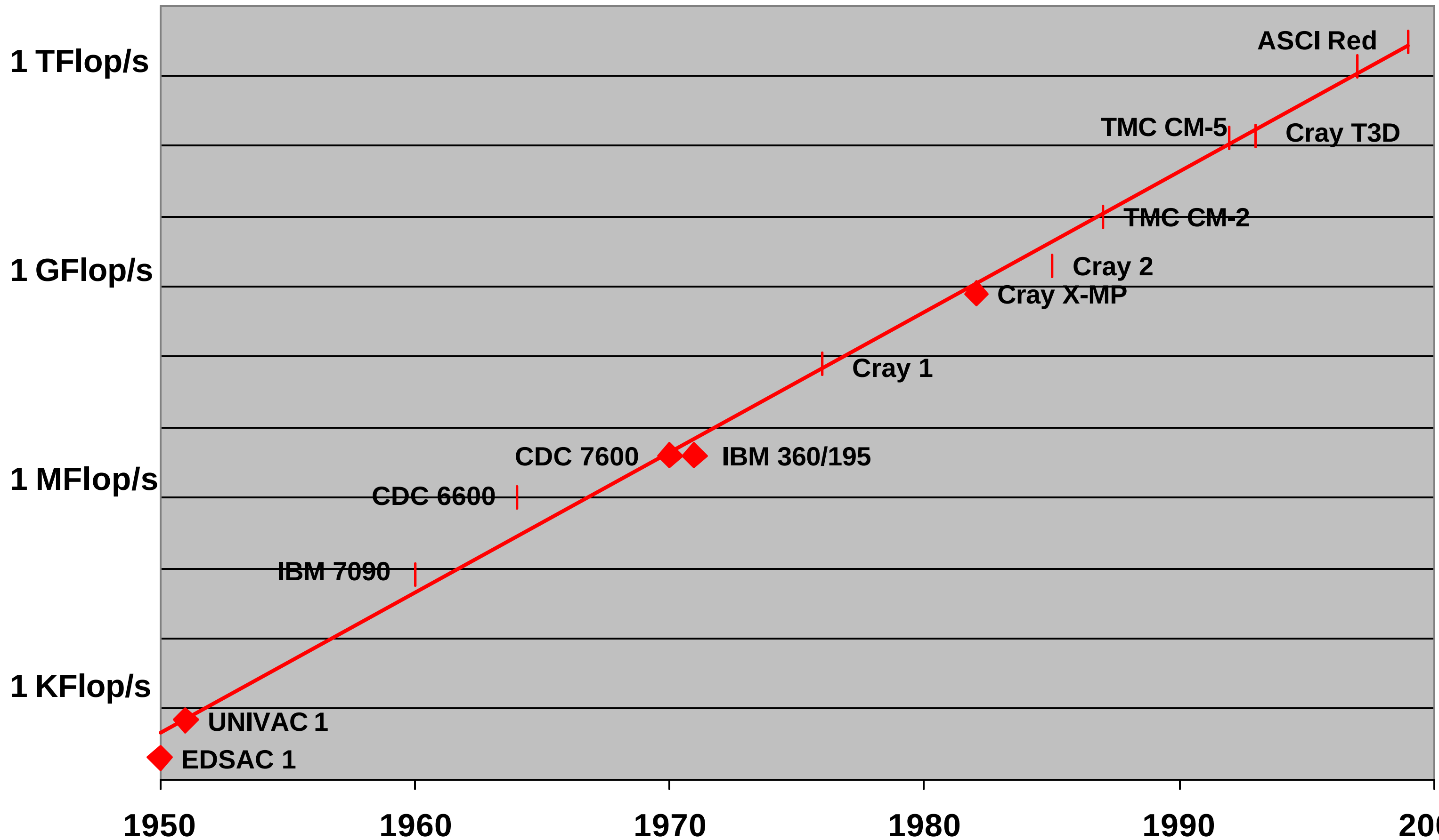
Livingston Plot: 10x energy increase every 6-8 years since 1930s





Panofsky and Breidenbach, Rev. Mod. Phys. 71, #2 (1999)

Moore's Law for HPC Peak Performance: 100x performance every decade



Intermission

IBM 1403 Printer (1964)

TeraFLOP systems are available now.

Why do we need them?

Are we ready to use them?

What are we doing with them?

Q: Why do we need terascale computing?
A: Design of Next-Generation Machines

- High accuracy requirements
  - —Design of 3D electromagnetic components
    - frequency accuracy to 1:10000
- Large-scale requirements
  - —Designing 3D electromagnetic components
    - system-scale modeling
  - —Modeling 3D intense beam dynamics
    - Halos, beam-beam effects, circular machines
  - —Modeling 3D advanced accelerator concepts
    - laser- and plasma-based accelerators
- More physics
  - —collisions, multi-species, surface effects, ionization, CSR, wakes,…

Q. Are we ready to use HPC systems?
A: Yes

1990+

Parallel Beam Dynamics (LANL)

Parallel Electromagnetics (SLAC)

1997

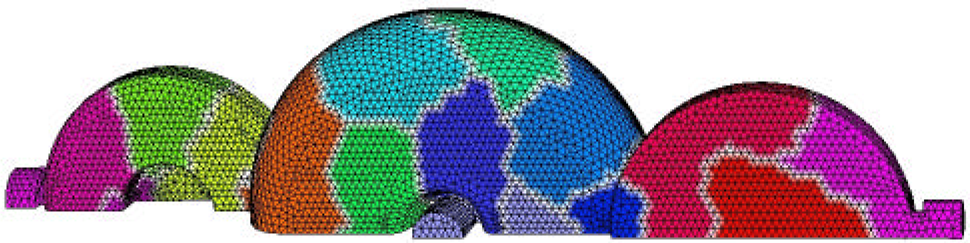
DOE/HPCC Grand Challenge (LANL, SLAC, UCLA, Stanford, ACL, NERSC)

2000

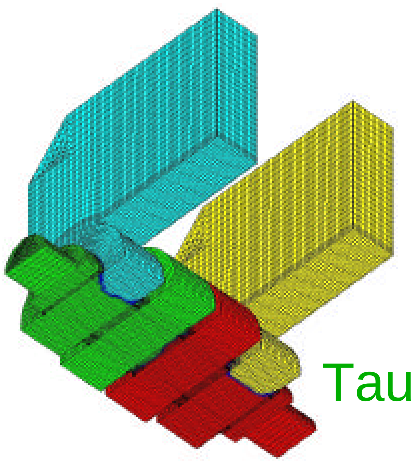
DOE/HENP extension (LANL, SLAC, LBNL, FNAL, BNL, Jlab, Stanford, UCLA, ACL, NERSC)

SciDAC project: <u>Advanced Computing for 21st Century Accelerator Science and Technology</u>

DOE Grand Challenge In Computational Accelerator Physics

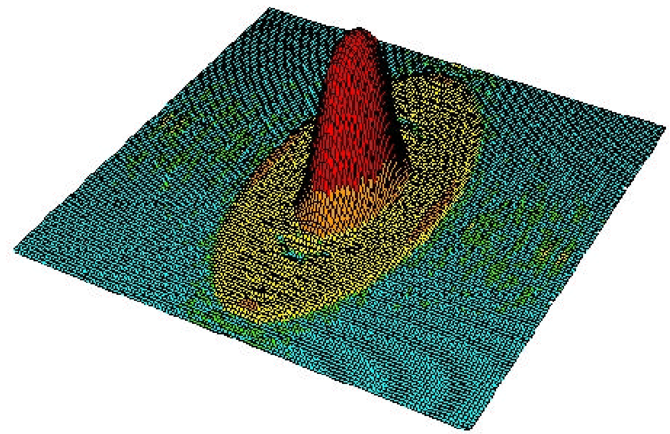


Omega3P: eigenmode



Tau3P: time-domain EM

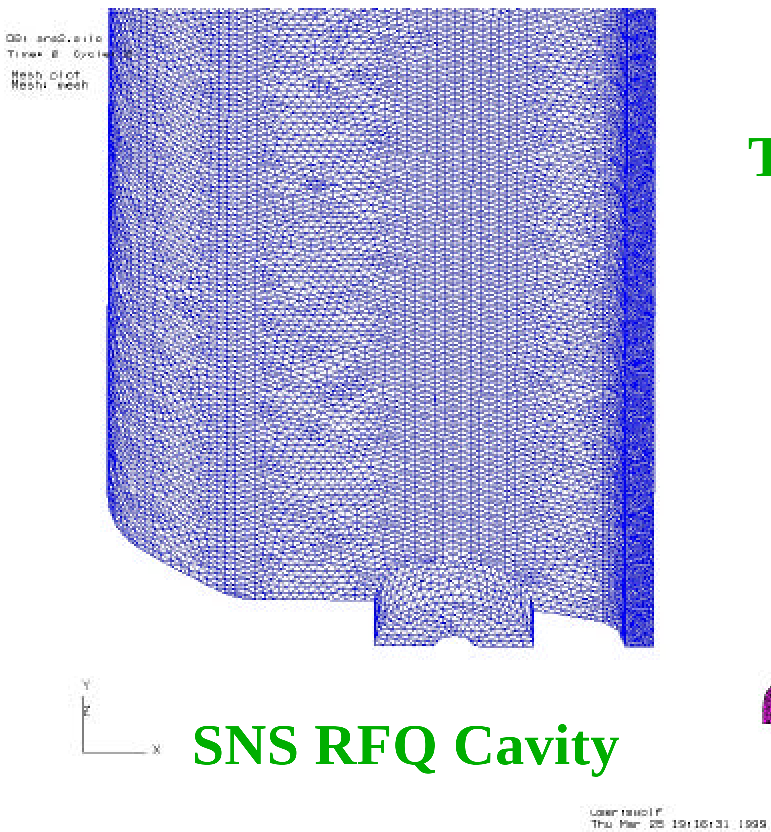
3 parallel application codes



IMPACT: Vlasov/Poisson

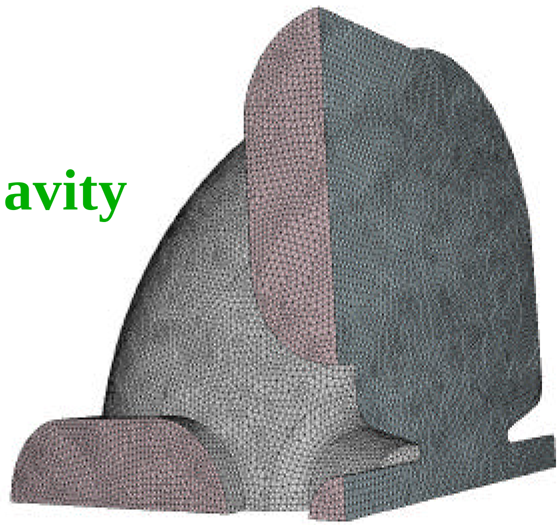
New capability has enabled simulations ***3-4 orders of magnitude*** greater than previously possible

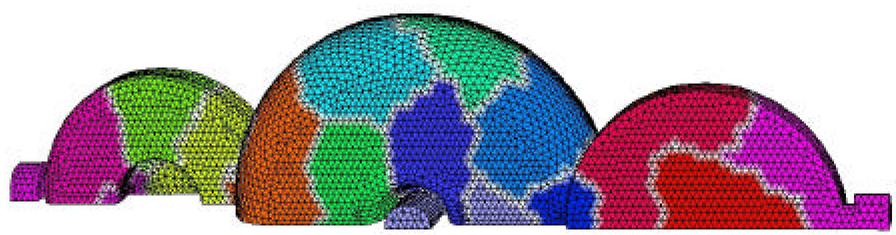
High Resolution Electromagneteic Modeling for Several Major Projects



SNS RFQ Cavity

TRISPAL Cavity

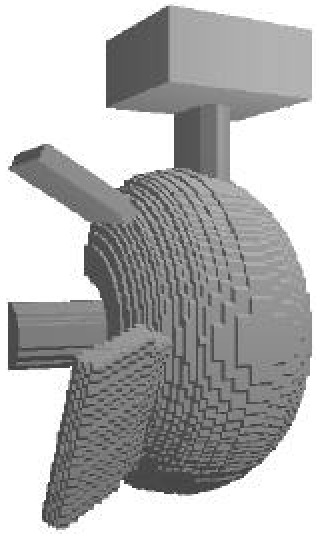




APT CCL Cavity

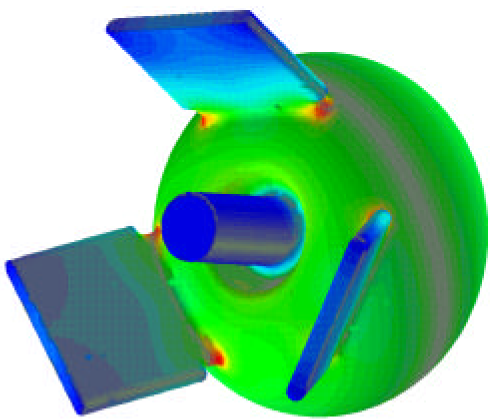
Mesh Refinement – Power Loss (Omega3P)

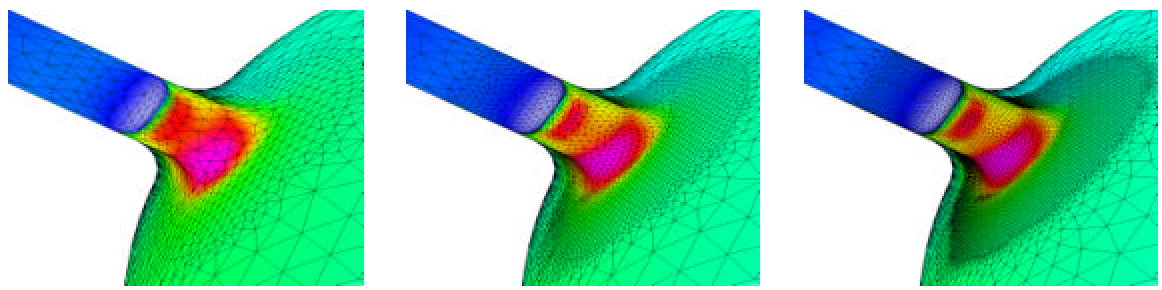
PEP-II Waveguide Damped RF cavity
- accurate wall loss distribution needed to guide cooling channel design



Structured Grid Model on single CPU

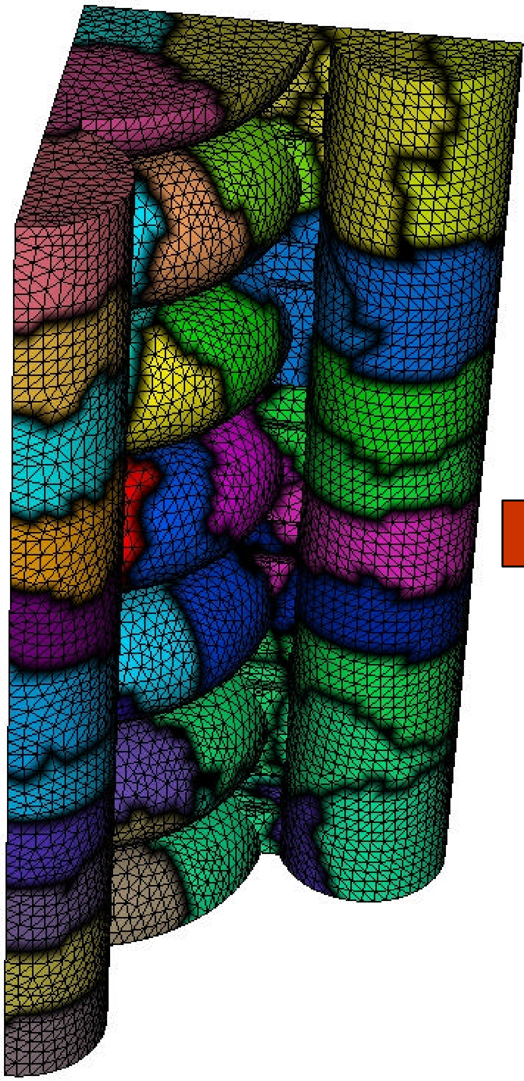
Parallel, Unstructured Grid Model – higher resolution





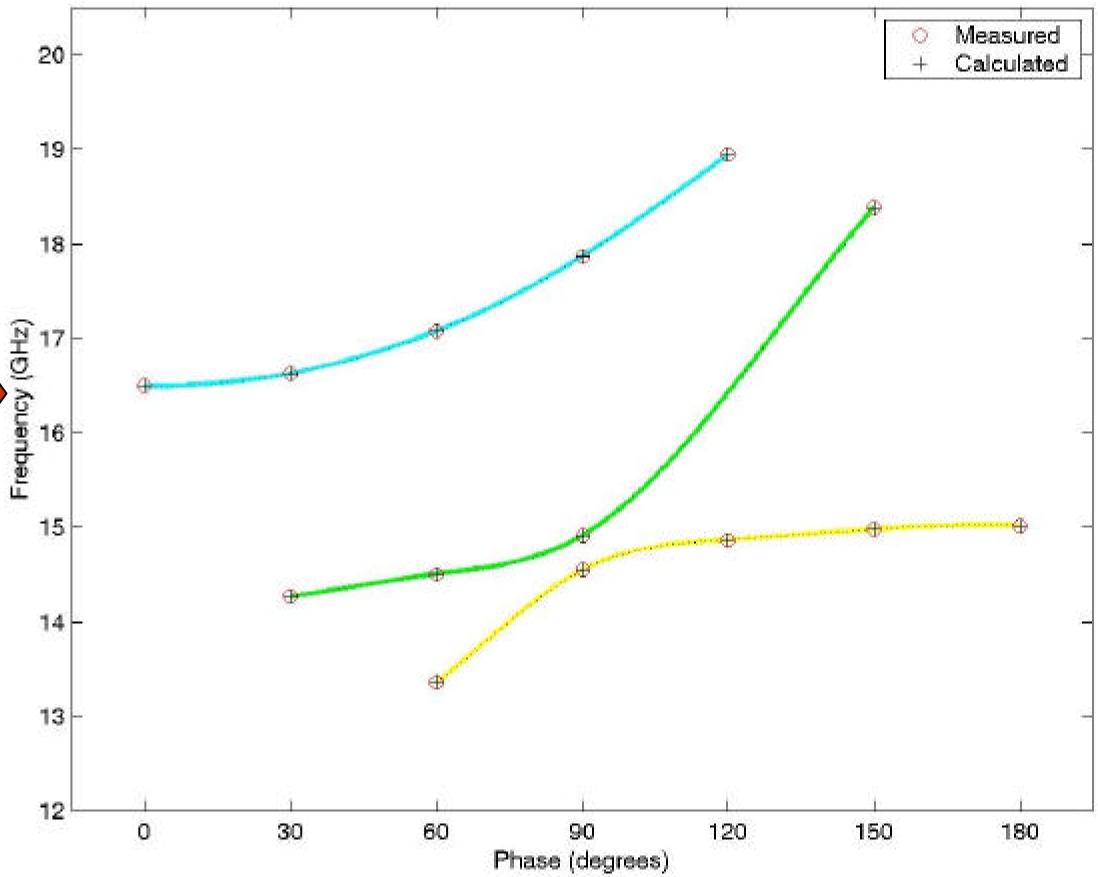
refined mesh size:	5 mm	2.5 mm	1.5mm
# elements :	23390	43555	106699
degrees of freedom:	142914	262162	642759
peak power density:	1.2811 MW/m^2	1.3909 MW/m^2	1.3959 MW/m^2

NLC RDDS Dipole Modes



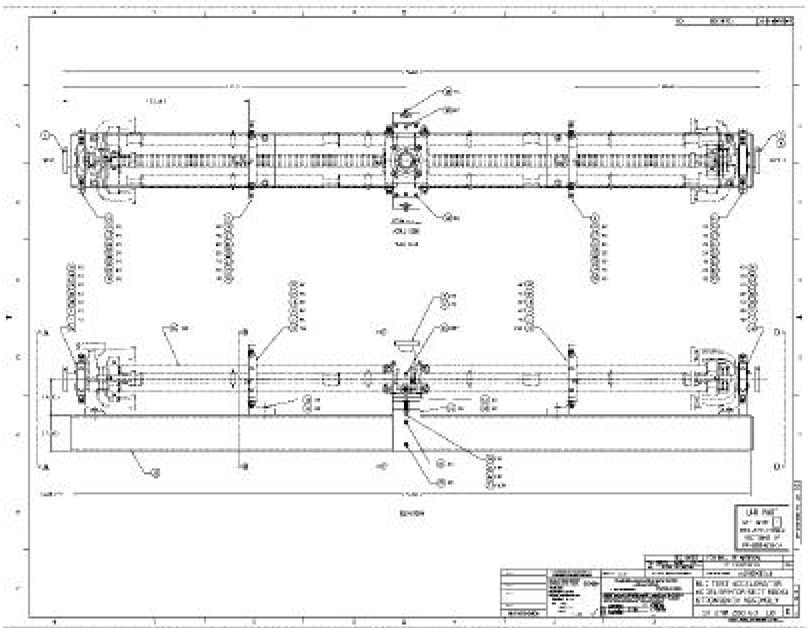
6 cell Stack

Lowest 3 dipole bands



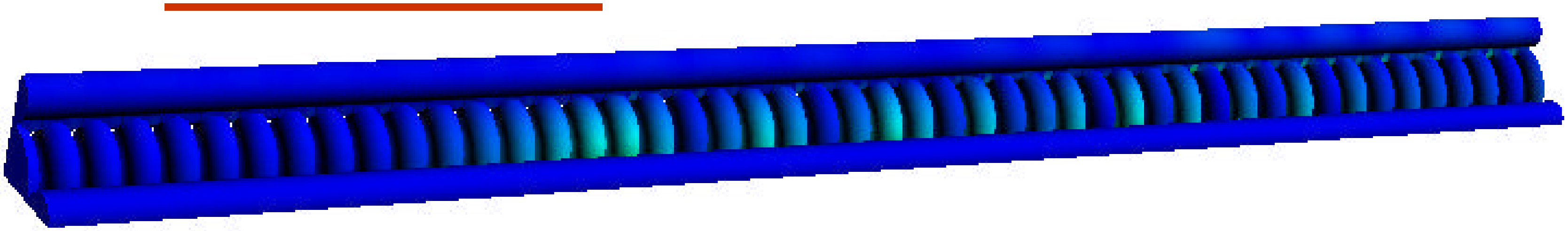
Toward Full Structure Simulation

RDDS 206-Cell Section



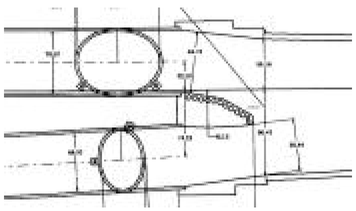
RDDS 47-Cell Stack

- Goal is to model entire RDDS section
- 47-cell stack is another step towards full structure simulation
- New low group structures are of comparable length, 53-83 cells
- Omega3P calculations become more challenging due to dense mode spectrum increasingly large matrix sizes (10's of millions of DOF's)

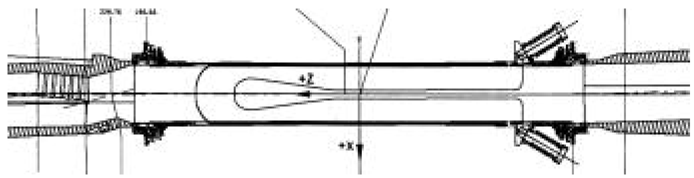


PEP II - IR Beamline Complex

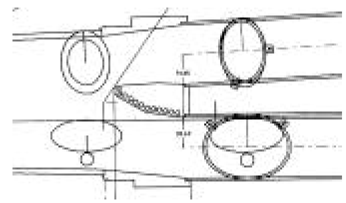
Left crotch | *Center beam pipe* | *Right crotch*

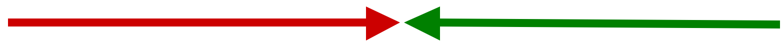


2.65 m



2.65 m

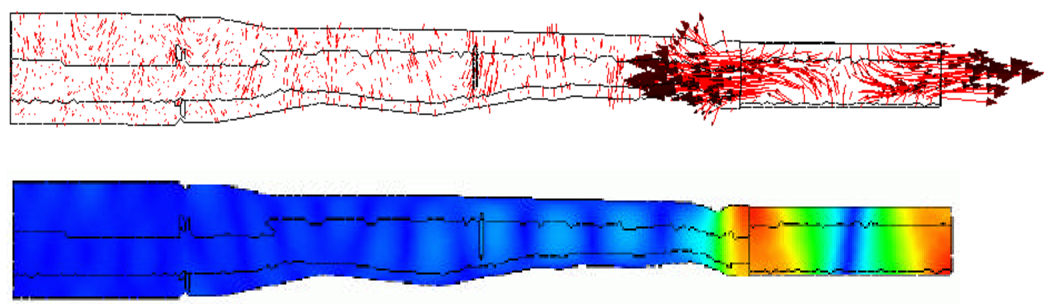




e_+ e_-

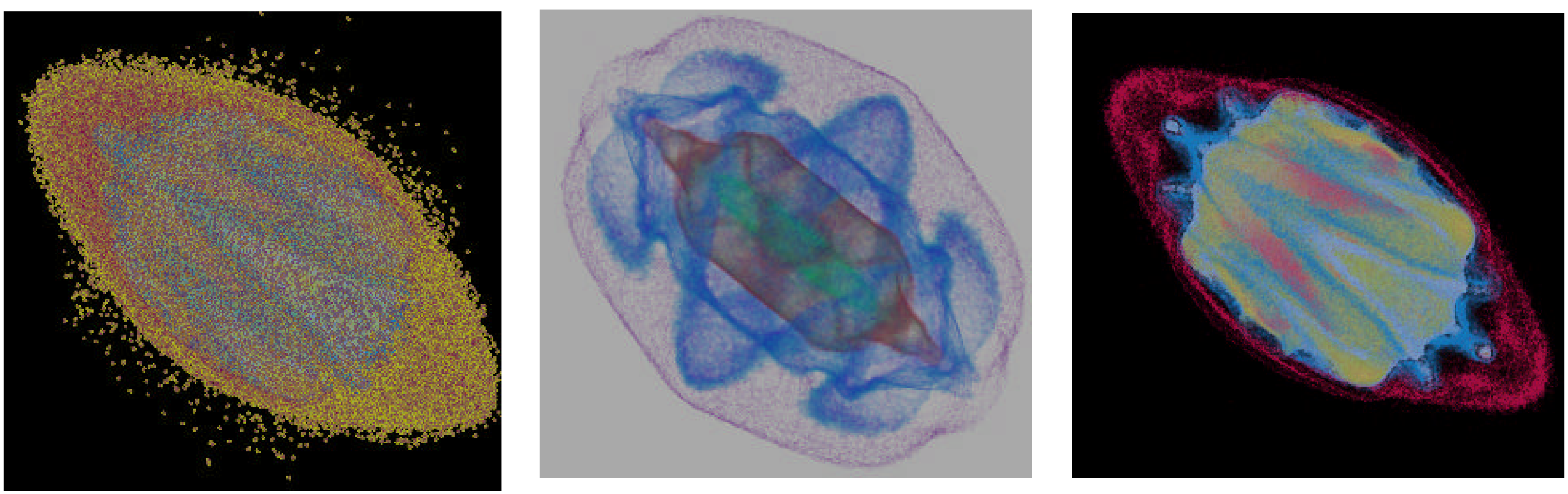
Identify localized modes to understand beam heating

Short section from IP



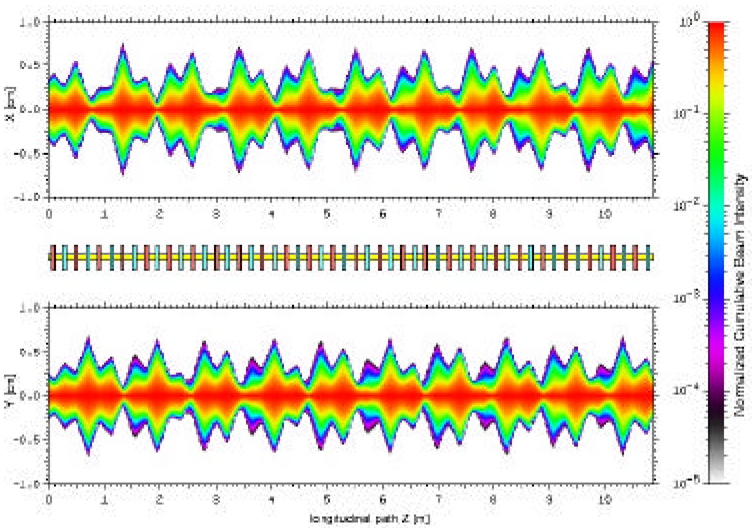
HPC Linac Modeling: 7 months reduced to 10 hours

- Beam dynamics problem size:
  —(128^3-512^3 grid points) x (~20 ptcls/point) = 40M-2B ptcls
- 2D linac simulations w/ 1M ptcls require 1 weekend on PC
  - 100Mp PC simulation, if possible, would take 7 months
- New 3D codes enable 100Mp runs in 10 hrs w/ 256 procs

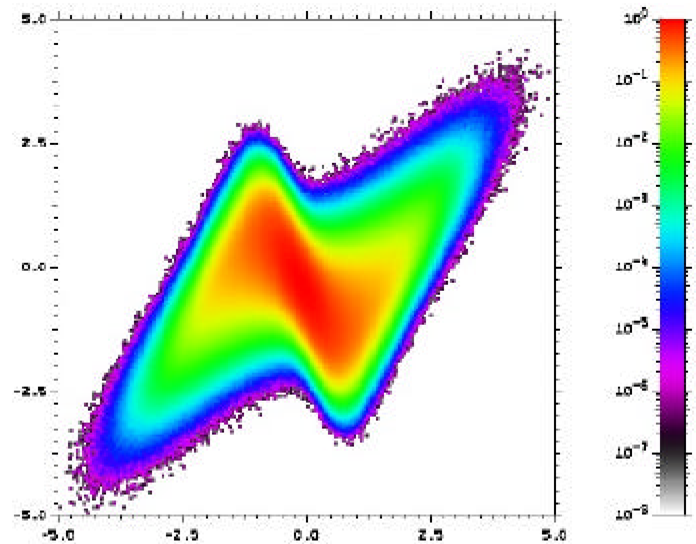


Beam Dynamics: Old vs. New Capability

- 1980s: 10K particle, 2D serial simulations
- Early 1990s: 10K-100K, 2D serial simulations
- 2000: 100M particle runs *routine* (5-10 hrs on 256 PEs); *more realistic model*



LEDA halo expt; 100M particles



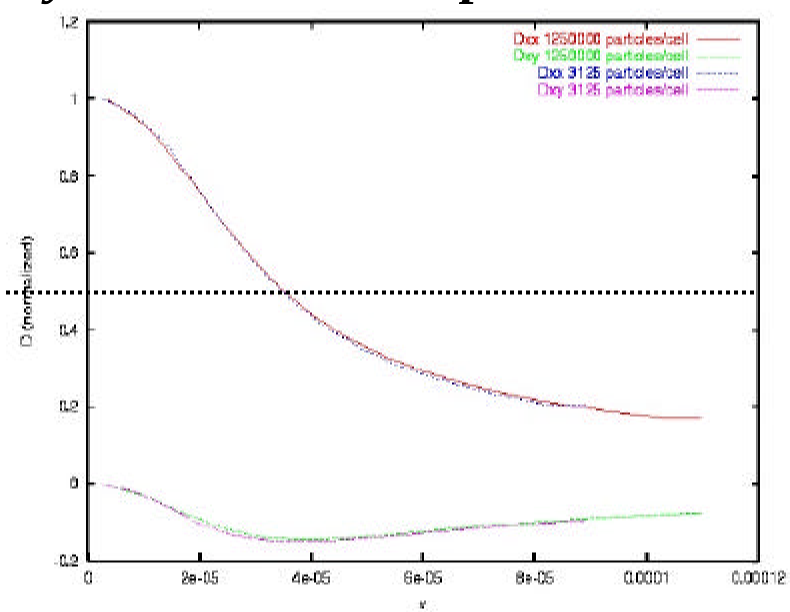
SNS linac; 500M particles

First-ever 3D Self-consistent Fokker-Planck Simulation (J. Qiang and S. Habib)

- Requires analog of 1000s of space-charge calculations/step
  - “…it would be completely impractical (in terms of # of particles, computation time, and statistical fluctuations) to actually compute [the Rosenbluth potentials] as multiple integrals” J.Math.Phys. 138 (1997).

FALSE. Feasibility demonstrated on parallel machines at NERSC and ACL

Self-Consistent Diffusion Coefficients



Spitzer approximation

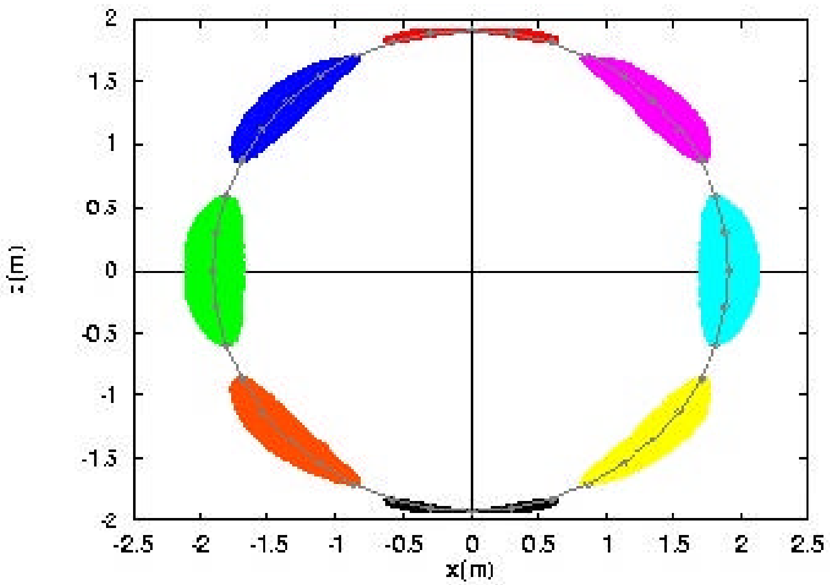
Previous approximate calculations performed w/out parallel computation were not self-consistent

High-Resolution Simulation of Intense Beams in Rings is a Major Challenge

- 100 to 1000 times more challenging than linac simulations
- Additional physics adds further complexity

x-z plots based on x-ϕ data from an s-code.

Data shown in a bend at different 8 times



We are approaching a situation where users will be able to "flip a switch" to turn space charge on/off in the major accelerator codes

Intermission

IBM 1403 Printer

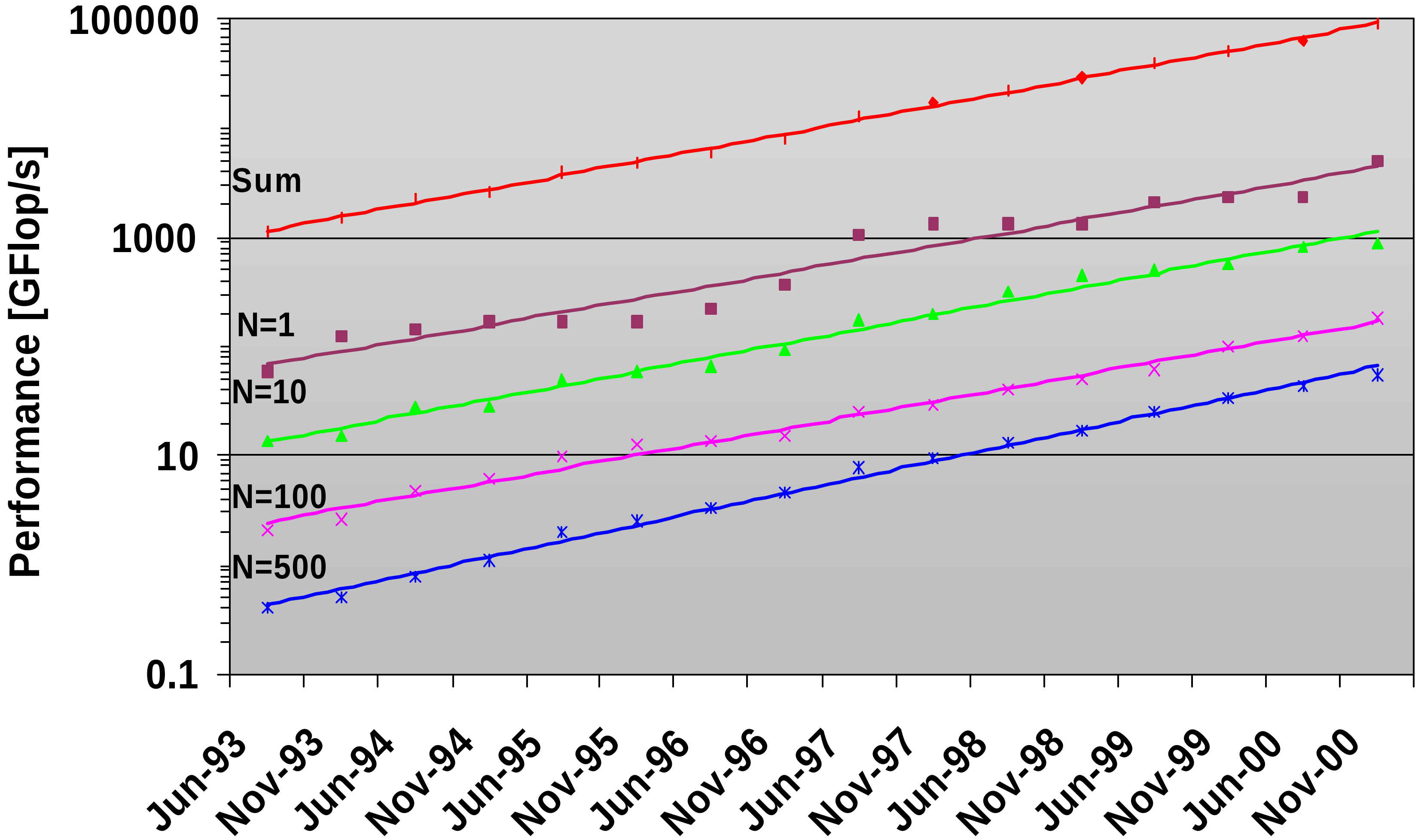
Berkeley Lab

What does the future hold for HPC* and for Accelerator Science?

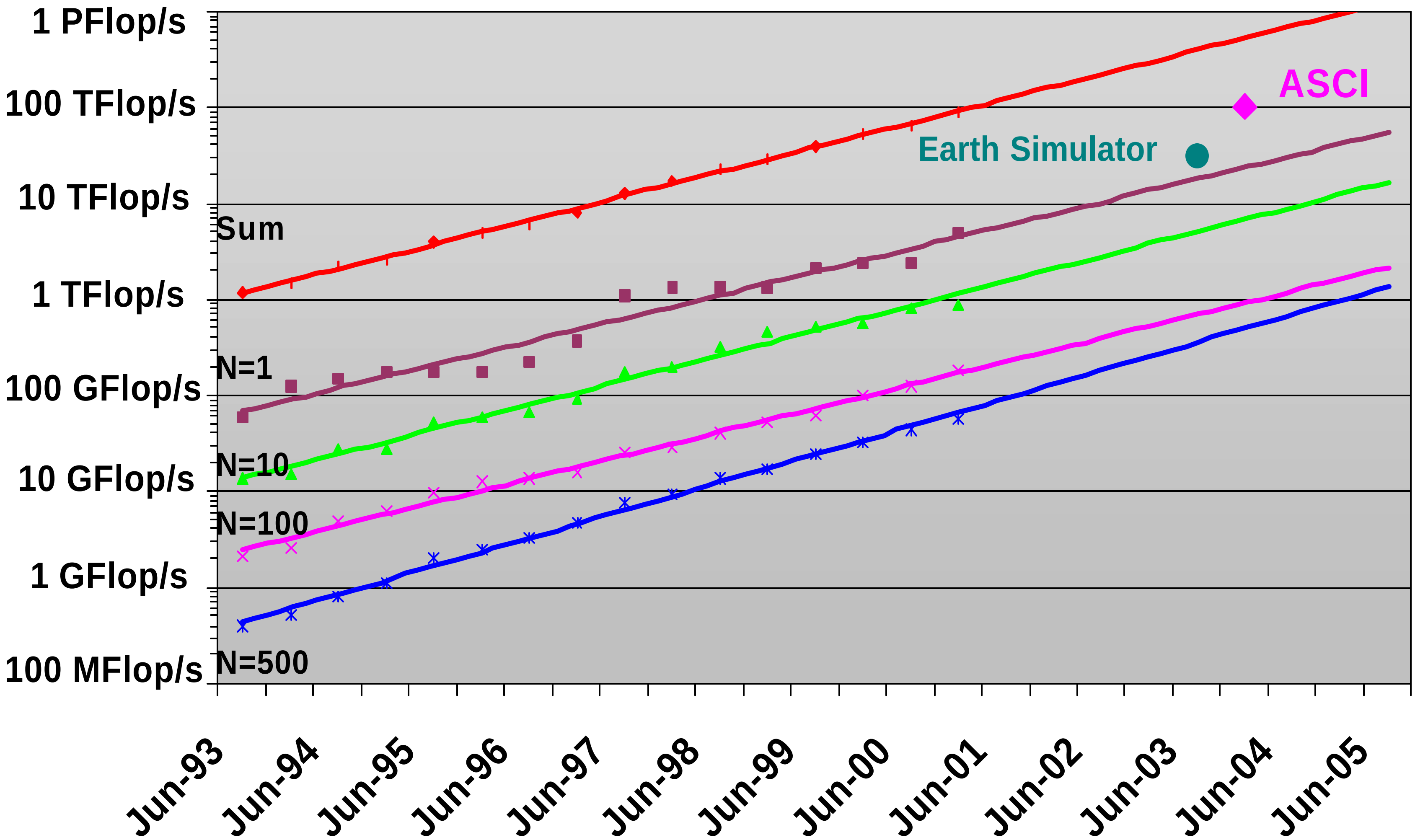
*David Bailey, NERSC

See also J. Dongarra and D. Walker, "The Quest for Petascale Computing," Computing in Science and Engineering, IEEE May/June 2001

Top500 List of Installed Supercomputers

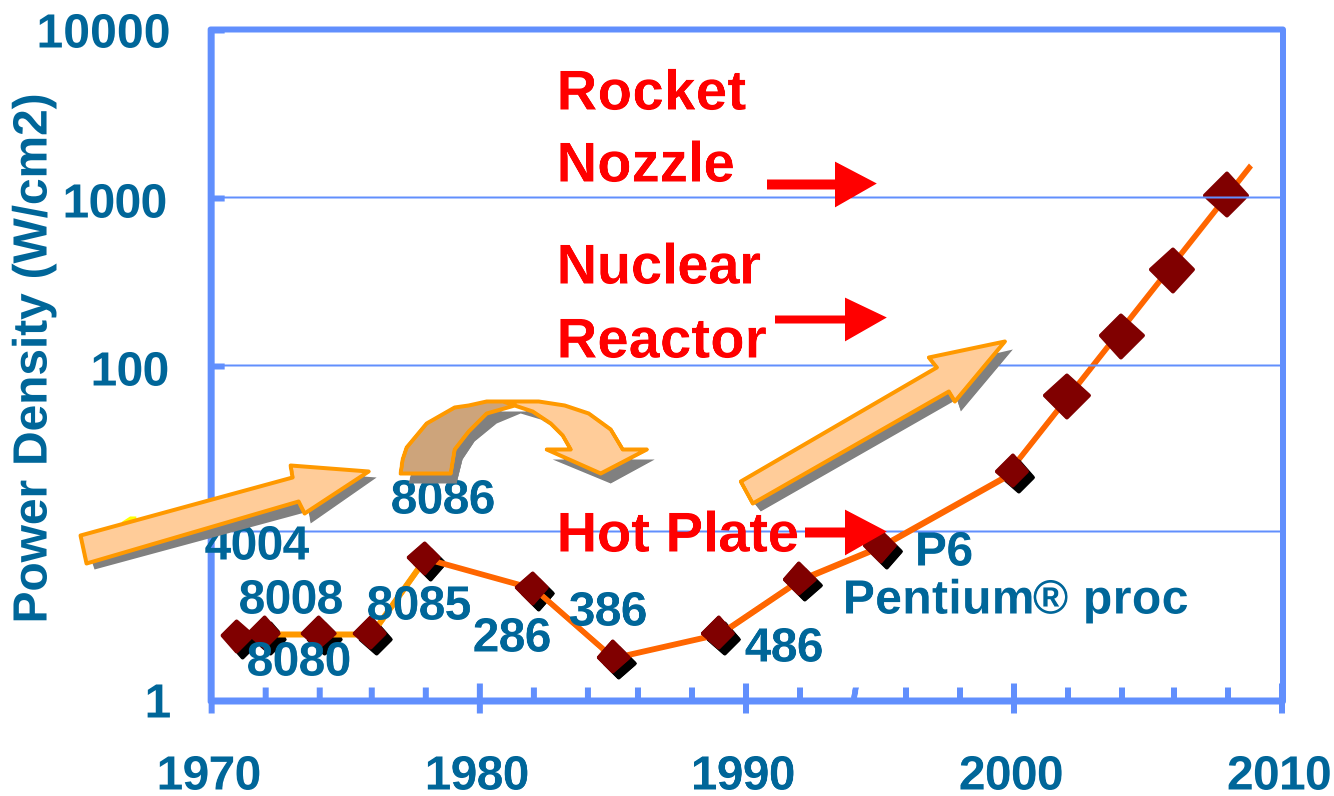


Top500 Extrapolation



Massive parallelism alone is not sufficient to reach the PetaFLOP regime

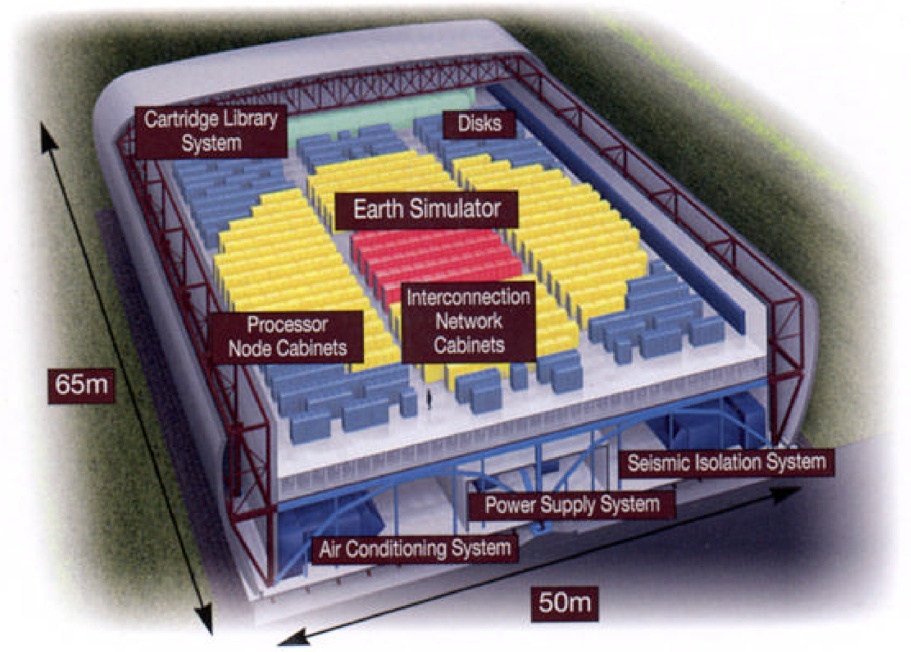
- Today: 10K-100K processors \$10B, 500 MW power
- Cannot simply wait for faster microprocessors

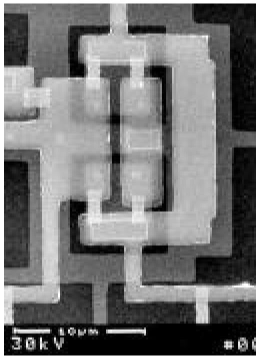


Shekhar Borkar, Intel

10-1000 TFLOP systems

- SMP clusters
  - —10 TFLOP @ LLNL
  - —30 TFLOP @ LANL
- Clusters with vector nodes
  - —Global Earth Simulator
- Special purpose machines
  - —IBM “Blue Gene”
  - —“Grape” system (N-body)
  - —Custom QCD systems
- New technologies/approaches
  - —Hybrid technology multi-thread (HTMT)



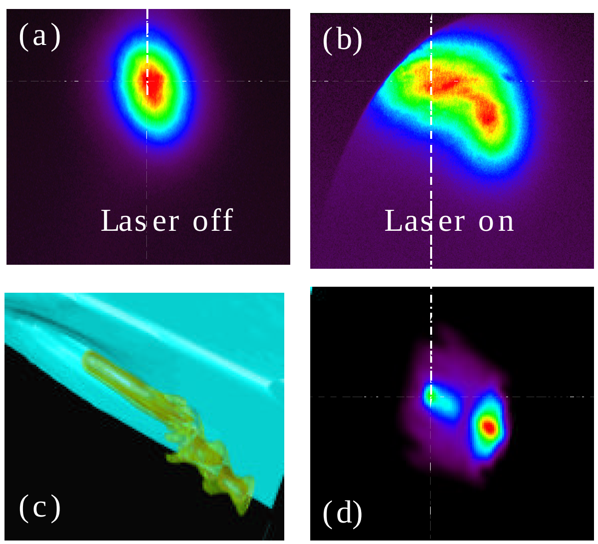


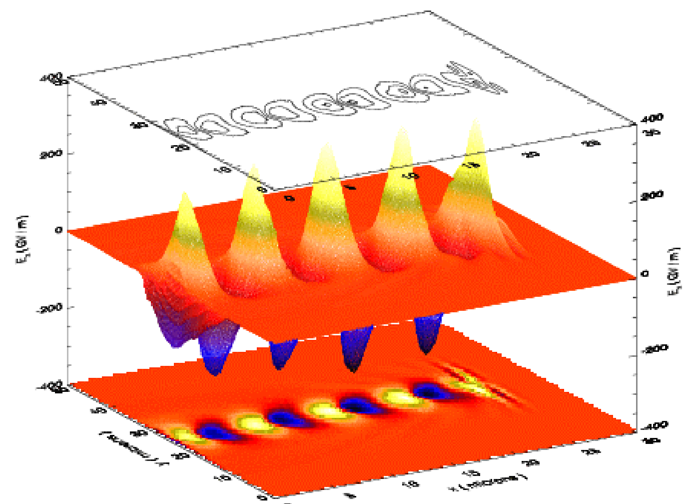
What must we do to maintain our pace?

- Smaller? Bigger?
- Higher performance
- Develop from technologies that have mass-appeal?
- Develop new technologies?

Experiment

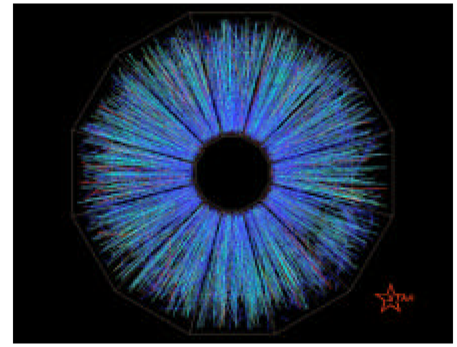
PIC Simulation

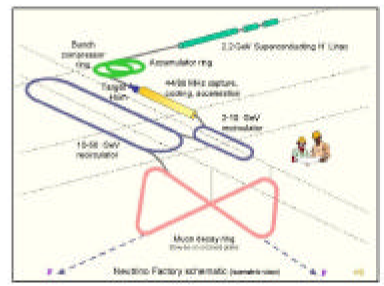


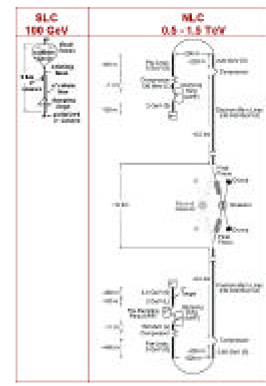


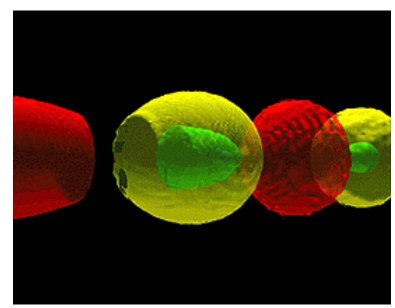
Summary: HPC will play a major role

- Present accelerators: Maximize investment by
  —optimizing performance
  —expanding operational envelopes
  —increasing reliability and availability
- Next-generation accelerators
  —facilitate important design decisions
  —feasibility studies
  —completion on schedule and within budget
- Accelerator science and technology
  —help develop new methods of acceleration
  —explore beams under extreme conditions

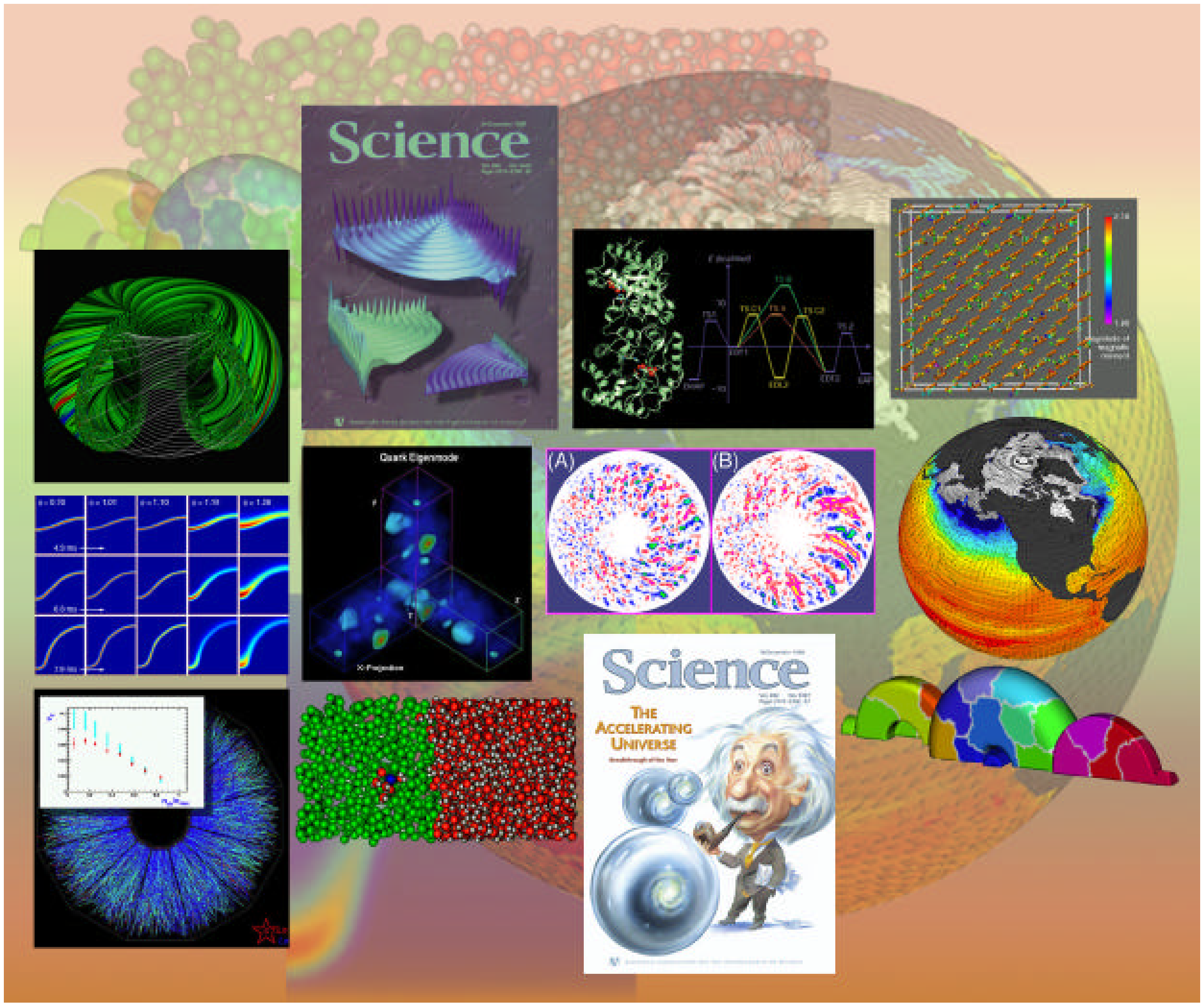








“… *computational science of scale* in which large teams attack fundamental problems in science and engineering that require massive calculations and have broad scientific and economic impacts”



HPC enables

Great Science in:

- Materials Science
- Climate
- Accelerator Physics
- Cosmology
- Molecular dynamics
- High Energy and Nuclear Physics
- Combustion
- Fusion
- Quantum Chemistry
- Biology
- much more…

Accelerator Science, like HPC, is an enabler of great science and greatly benefits society